Evidence-Based Policy and the Governance of Evidence in Indonesian Local Government: A Single-Case Study of Policy Capacity in Kudus Regency

Ismi Darmastuti[1]*, Ahyar Yuniawan[2], Aulia Vidya Almadana[3]

[1,2,3] Department of Management, Diponegoro University, Semarang, Indonesia

ARTICLE INFO

Article history:
Received: August 21, 2025
Revised: November 10, 2025
Accepted: December 1, 2025

JEL Classification:
H75, H83, I28,O38

DOI:
10.14414/jebav.v28i3.5732

Keywords:
decentralization, evidence-based policy good governance of the evidence; policy capacity: a single-case study.

Corresponding Author:
ismidarmastuti.febundip@gmail.com

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ABSTRACT

*This study explores how the Kudus Regency in Central Java organizes evidence-based policy (EBP) demand and supply, presenting a context-sensitive model for local governance. Conducted through a two-stage qualitative case study (April 2022–February 2023), it involved 16 interviews, three focus groups, and documentary analysis with stakeholders from Bappeda, OPDs, Diskominfo, academia, and civil society. Stage 2 refined the EBP model via participatory workshops with 22 participants, using NVivo for thematic coding (κ = 0.81) and member-checking. Findings reveal EBP as an input-formulation-implementation-outcome cycle constrained by three intersecting deficits: analytical, operational (especially metadata standardization and interoperability), and political. The "3C principle" (cooperation, communication, coordination) emerged as a meso-level mechanism bridging individual analytical efforts to inter-organizational outcomes. While the single-case, perceptual evidence limits causal claims, the study proposes an EBP formulation model integrating *Satu Data Indonesia* architecture with deliberative forums like Musrenbang, alongside a diagnostic template to identify binding constraints in local evidence cycles.*

ABSTRAK

Studi ini mengkaji bagaimana Kabupaten Kudus, Jawa Tengah, mengorganisasikan permintaan dan penyediaan bukti untuk mendukung kebijakan berbasis bukti (evidence-based policy, EBP) dalam konteks pemerintahan daerah, dua dekade setelah desentralisasi di Indonesia. Penelitian ini menggunakan studi kasus kualitatif yang dilaksanakan dalam dua tahap. Tahap pertama (April–September 2022) mencakup wawancara mendalam dengan 16 informan, tiga diskusi kelompok terfokus, serta analisis dokumen. Informan dipilih dari Bappeda, OPD, Diskominfo, kalangan akademisi, dan masyarakat sipil. Tahap kedua (Oktober 2022–Februari 2023) menguji model EBP melalui dua lokakarya partisipatif yang diikuti oleh 22 peserta. Data dianalisis secara tematik, dilengkapi dengan member checking, dan diverifikasi melalui triangulasi lintas sumber. Hasil penelitian menunjukkan bahwa EBP di Kudus berlangsung melalui siklus masukan–perumusan–implementasi–hasil, tetapi masih dibatasi oleh tiga defisit kapasitas, yaitu kapasitas analitis, operasional, dan politik.

1. INTRODUCTION

Indonesia's transition from a centralized to a decentralized system of government, initiated by Law No. 22 of 1999 and consolidated through Law No. 23 of 2014, transferred substantial authority, fiscal resources, and political accountability to provinces, regencies, and cities, making the local government the principal arena in which citizens encounter the state. Two decades later, the developmental dividends of this transfer have proven geographically uneven and contingent on local capacity rather than guaranteed by formal devolution (Shoesmith, Franklin, & Hidayat, 2020; Sutiyo & Maharjan, 2017; Hummel & Kusumasari, 2025). How local governments convert devolved authority into effective, citizen-responsive policy—and what capacities this process requires—therefore remains both empirically pressing and theoretically central to the literature.

Evidence-based policy (EBP) is an influential approach that holds that public decisions should be grounded in rigorous, systematically generated evidence rather than habit, lobbying, or ideology (Davies, Nutley, & Smith, 2000; Nutley, Walter, & Davies, 2007; Head, 2016). EBP informs the OECD's (2020) governance agenda and Indonesia's reform programme, particularly through Presidential Regulation No. 39 of 2019 on *Satu Data Indonesia* (One Data Indonesia) and the e-government framework established under Presidential Regulation No. 95 of 2018. These instruments have also been adopted at the regional level. For example, Kudus Regency issued Regent Regulation No. 41 of 2021 on the Implementation of One Regional Data to standardize data governance and support evidence-informed planning.

Despite this institutional momentum, the available evidence suggests that the substantive use of evidence in Indonesian local governments remains limited. Sutmuller and Setiono's (2011) baseline study of 21 local governments found that only a small proportion routinely formulated policies based on empirical analysis; many replicated the strategies of neighboring regions or maintained existing policies without evaluation. More recent scholarship reaches similar conclusions: policy analytical capacity in Indonesian subnational bureaucracies is uneven and tends to be concentrated in a small number of well-resourced units (Newman, Patmisari, & Widianingsih, 2022). The implementation of *Satu Data Indonesia* has been hindered by sectoral silos, fragmented metadata, and insufficient interoperability among agencies (Bernot et al., 2024). Taken together, these findings point to a persistent gap between the formal architecture of EBP and its de facto practice in regional bureaucracies.

Two interrelated problems account for this gap in the literature. First, early literature on EBP tended to treat evidence use primarily as a technical matter, overlooking the political and institutional conditions under which evidence enters the policy cycle (Cairney, 2016, 2022; Parkhurst, 2017). Second, both academic and policy literature on Indonesian local government have paid relatively little attention to how policy capacity—analytical, operational, and political—interacts with the demand for and supply of evidence at the regency level, where most public services are delivered (Wu, Ramesh, & Howlett, 2015; Howlett, 2015; Mukherjee, Coban, & Bali, 2021; Aliyu, Soeparman, & Hartono, 2022). Consequently, normative prescriptions concerning EBP often fail to engage with the concrete organizational, informational, and political conditions facing Indonesian regents and their planning agencies. This study addresses this gap by developing a context-sensitive EBP formulation model for Indonesian local governments and refining it using evidence from a single regency. Two research questions guided the analysis:

(RQ1) How do local governments collect, interpret, and use evidence at the input, formulation, implementation, and outcome stages of the policy cycle?

(RQ2) Which capacity, governance, and political conditions enable or constrain EBP, and how can these conditions be incorporated into a feasible formulation model?

Two clarifications regarding the epistemic status of the following discussion are necessary at the outset, as they inform the interpretation of each claim. First, the research design was qualitative, abductive, and interpretive. It seeks to describe and theorize a mechanism rather

than estimate an effect, and it does not test hypotheses in a statistical sense. Second, the outcome evidence presented in Section 4.4 consists of assessments by actors involved in the policy process concerning their domains. The use of formulations such as "informants associated," "was reported as, " or "was perceived to" is deliberate; no causal claim is intended or supported. This distinction is consistently maintained from the abstract to the conclusion.

This study makes three contributions, each calibrated to the evidence a single case can support. **Theoretically**, it brings the policy-capacity literature into sustained dialogue with Parkhurst's (2017) good governance of evidence and the collaborative-governance literature on inter-organizational coordination (Ansell & Gash, 2008; Emerson, Nabatchi, & Balogh, 2012), treating evidence not only as an epistemic resource but also as a politically governed flow. This integration is not entirely novel conceptually—recent scholarship on the governance of evidence already moves in this direction (Weible & Cairney, 2021; Christensen, 2025)—but it has not previously been operationalized as an analytical instrument at the Indonesian district level. **Empirically**, the article provides one of the few detailed accounts of how an Indonesian regency operationalizes the post-2019 *Satu Data Indonesia* architecture in day-to-day policy settings. It complements, rather than supersedes, Newman et al.'s (2022) provincial-level analysis of West Java by shifting the unit of analysis to the regency, where most basic services are delivered and analytical infrastructure is thinnest. **Practically**, the proposed input–formulation–implementation–outcome model provides practitioners with a diagnostic template for identifying where their own evidence cycle breaks down.
The remainder of this paper is organized as follows. Section 2 reviews debates concerning EBP, policy capacity, collaborative governance, and the politics of evidence, and develops the analytical framework and four orienting propositions. Section 3 describes the research design, case selection, sampling, saturation protocol, and analytical strategy. Section 4 presents the findings from Kudus Regency across the four stages of the EBP cycle. Section 5 discusses the findings in relation to the literature, develops the proposed model, considers alternative interpretations and threats to construct validity, and specifies the boundary conditions for transferability. Section 6 concludes.

2. LITERATURE REVIEW AND ANALYTICAL FRAMEWORK

Four terms recur throughout the article and are used consistently in the sense defined here. *Evidence-based policy (EBP)* refers to the normative program of grounding public decisions in systematically generated evidence. Good *governance of evidence* refers to Parkhurst's (2017) reframing of the program as a question of how the use of evidence is institutionally governed and assessed against the criteria of credibility, salience, and legitimacy. *Policy capacity* refers to the Wu–Ramesh–Howlett (2015) triad of analytical, operational, and political competencies across the individual, organizational, and systemic levels. The *EBP cycle* refers to the four-stage sequence—input, formulation, implementation, and outcome—used as the analytical framework. Sections 2.1 and 2.2 develop the first two terms and the third term, respectively. Section 2.3 situates them within the Indonesian institutional setting, and Section 2.4 integrates them into the analytical framework and its propositions.

2.1 From evidence-based policy to the good governance of evidence

EBP emerged in the late 1990s as a transposition of evidence-based medicine into public policy, championed by the United Kingdom's modernizing government agenda and subsequently disseminated through international organizations (Davies et al., 2000; Sanderson, 2002). At its core, EBP holds that policymakers should integrate experience and judgment with the best available external evidence derived from systematic research (Nutley et al., 2007). Early proponents distinguished EBP from *opinion-based policy*, in which decisions

rest primarily on expert judgment, and from purely *ideological policy*, in which they are grounded solely in normative commitments (Head, 2016).

Two decades of empirical research have complicated normative ideals. Oliver et al.'s (2014) systematic review of 145 studies identified a recurring set of barriers—limited contact between researchers and policymakers, time pressures, opaque research findings, and political incentives that reward symbolic rather than substantive use of evidence—that impede the rationalist model of evidence use in policy. Cairney (2016, 2022) argues that evidence-based policymaking is best understood as an aspiration rather than a description of practice: policymaking takes place in complex, decentralised environments where bounded rationality, bargaining among multiple actors, and competing problem framings preclude the linear use of evidence. Building on this critique, Parkhurst (2017) reframes the question from *"How can policy be more evidence-based?"* to *"How can the use of evidence be governed well*?" His concept of good governance of evidence specifies that legitimate evidence use requires not only credible and salient evidence but also institutions that are representative of and accountable to the affected populations.

More recent contributions further extend this shift from epistemics to institutions. Christensen (2025) demonstrates that organizational structures—including the design of in-house expertise, advisory committees, and analytical units—have a measurable, independent effect on evidence uptake. Weible and Cairney (2021) draw on policy theory to recommend integrating deliberative engagement with technical analysis. Höchtl, Parycek, and Schöllhammer (2016) and Giest et al. (2025) document how digital and data-driven transformations—including open-data portals, interoperable platforms, and algorithmic analytics—are reshaping the evidence ecosystem of contemporary public administration. Dunlop and Radaelli (2018) situate the resulting feedback processes within the broader literature on policy learning. Taken together, these contributions suggest that EBP is best conceptualized as a multi-stage, politically embedded socio-technical process rather than a unidirectional transmission of research into decision-making.

If EBP is an institutionally governed process rather than a technical transmission, the immediate question is which institutional competencies such governance requires and where these competencies reside. The policy capacity literature reviewed below provides the most comprehensive answer to this question.

2.2 Policy capacity, collaboration and local government

Wu, Ramesh, and Howlett (2015, 2018) define policy capacity as the set of skills and resources required to perform policy functions, organized across three substantive competencies—analytical, operational, and political—and three resource levels—individual, organizational, and systemic. Analytical capacity encompasses the knowledge, methods, and information needed to diagnose problems and design responses. Operational capacity refers to the human, financial, and organizational resources required to coordinate and implement decisions. Political capacity comprises the legitimacy, leadership, and stakeholder engagement needed to mobilize support. Howlett (2015) offers an important refinement for the present study: analytical capacity depends not only on supply (trained analysts) but also on demand (decision-makers who require analysis); consequently, investment in the former may yield little when the latter is absent.

Howlett and Ramesh (2016) argue that governance failures typically stem from critical capacity deficits—often described as "Achilles' heels"—that compromise the entire policy cycle. Mukherjee, Coban and Bali (2021) further demonstrate that the three competences are complementary: strong analytical capacity is ineffective without operational delivery systems, while operational systems may produce arbitrary outcomes without political legitimacy. Newman, Cherney and Head (2017) directly link this framework to evidence use, showing in Australian survey data that policy analytical capacity is the strongest correlate of public servants' self-reported evidence use.

A related body of literature approaches the same problem from the opposite direction, asking not what competences an organization possesses but how organizations with different competences are brought into productive relationships. Ansell and Gash (2008) identify face-to-face dialogue, trust-building, and the incremental accumulation of shared understanding as the conditions under which multi-actor collaboration produces decisions rather than deadlocks. Emerson, Nabatchi, and Balogh (2012) formalized these conditions in an integrative framework in which principled engagement, shared motivation, and joint capacity for action constitute the dynamics of collaborative governance. This literature is relevant to the present article because the meso-level mechanism reported in Section 4.2—what informants call the *3C principle*—is, conceptually, a vernacular articulation of precisely these collaborative dynamics, operating between agencies within a single local government rather than between state and nonstate actors. Interpreting the inductive findings in light of this established literature, rather than presenting them as sui generis, is one of the analytical contributions made in Section 5.

Empirical applications in Indonesia and the wider Southeast Asian region have expanded rapidly. Drawing on panel data from Indonesian regencies, Aliyu, Soeparman and Hartono (2022) demonstrate that the interaction among analytical, operational, and political capacities—rather than any individual capacity in isolation—predicts public service delivery outcomes. Saguin, Ramesh and Howlett (2018) report similar findings in the Philippines. Newman, Patmisari and Widianingsih (2022) examine policy analytical work in the provincial government of West Java and identify substantial evidence of policy analysis activity recognizable to Western scholarship, alongside persistent challenges involving fragmentation, weak demand, and limited institutionalization. These studies reinforce a key implication: a viable EBP model for Indonesian local government must be designed to reflect the actual configuration of capacities available at the regency level, rather than imported wholesale from comparator jurisdictions.

However, capacity and collaboration are invariably exercised within a specific institutional ecology. Two features of the Indonesian context shape the movement of evidence within local government: a mandated deliberative planning forum and a recently centralized data architecture. Section 2.3 examines each of these features.

2.3 Decentralisation, deliberation and the Indonesian setting

Indonesian decentralization has created a distinctive institutional ecology in which evidence and deliberation must coexist. The *Musyawarah Perencanaan Pembangunan* (Musrenbang), or Development Planning Deliberation Forum, is the primary mandated forum through which local governments are expected to consult citizens and consolidate development priorities (Anindito, Sagala, & Tarigan, 2022). Empirical studies, however, have criticized Musrenbang processes for fostering procedural and symbolic participation, whereby attendance is high but substantive influence over decisions remains limited (Sutiyo & Maharjan, 2017). Fossati (2016) demonstrates that local responsiveness to poverty in Indonesian districts depends on competitive electoral politics rather than planning procedures alone. These findings underscore the importance of the political dimension of EBP emphasized by Cairney and Parkhurst: technocratic data work cannot replace legitimate deliberative forums and must instead be complemented by them.

On the data supply side, the post-2019 *Satu Data Indonesia* architecture has begun to standardize data governance through the Bappenas Secretariat, while regions have established data custodian (*walidata*) functions within their Diskominfo and Bappeda offices (Bernot et al., 2024). Simultaneously, evaluations indicate persistent challenges: in many regencies, data-sharing mechanisms remain poorly managed; technical and human resource constraints limit interoperability; and OPDs frequently rely on intuition or precedent rather than statistical analysis. Comparative research on digitalized local administration suggests that these constraints are not unique to Indonesia. For example, Hirsch et al. (2022) find that

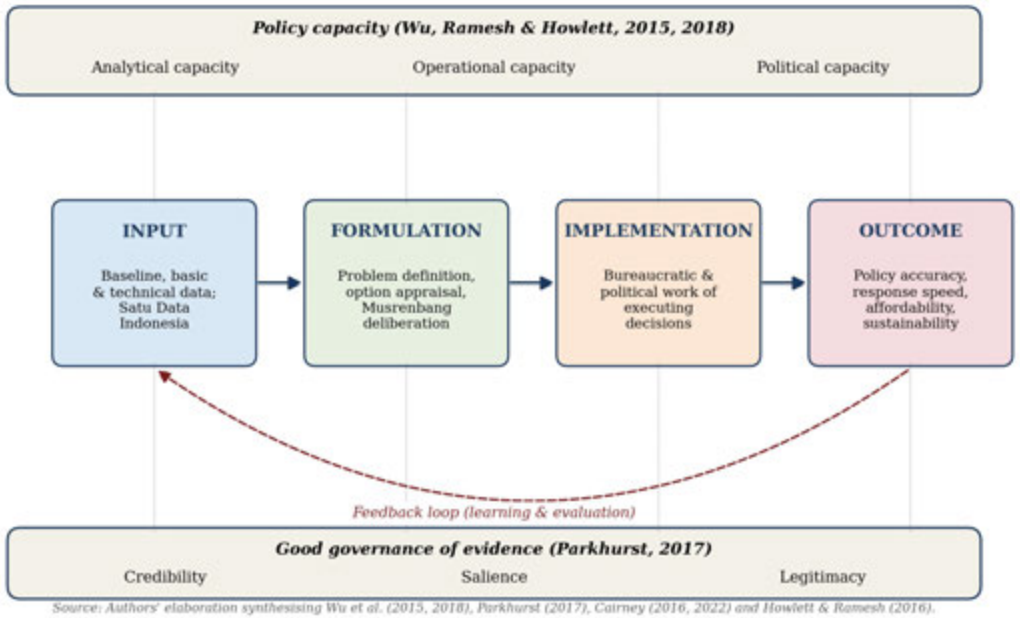
the effectiveness of digitalized communication in Central European local authorities depends on established organizational routines rather than technology. Indonesia's institutional framework is in place, but the day-to-day practices required to make it operational are still being developed.

Sections 2.1–2.3 thus identify three requirements for a workable account of local EBP: a stage-wise view of how evidence enters and moves through the policy process; an explicit treatment of the capacities that condition each stage; and attention to the arenas in which technical and political claims are reconciled. Section 2.4 combines these aspects.

2.4 Analytical framework and orienting propositions

Synthesizing these strands, the article adopts a four-stage analytical framework—input, formulation, implementation, and outcome—and examines each stage through the policy-capacity triad (analytical, operational, and political) and the criteria for good governance of evidence (credibility, salience, and legitimacy). Inputs comprise baseline, basic, and technical data, together with the institutional architecture of *Satu Data Indonesia*. Formulation encompasses problem definition, option appraisal, and decision-making by Bappeda, OPDs, and political principals, mediated by deliberative forums such as Musrenbang. Implementation refers to the bureaucratic and political processes involved in executing decisions. Outcomes comprise the effects on local government performance, operationalized in terms of policy accuracy, response speed, service affordability, and programme sustainability. The framework is summarized in Figure 1.

Figure 1. Analytical framework for evidence-based policy in Indonesian local government



Source: Authors' elaboration synthesising Wu et al. (2015, 2018), Parkhurst (2017), Cairney (2016, 2022) and Howlett & Ramesh (2016).

Figure 1. Analytical framework: an evidence-based policy cycle for Indonesian local government – input → formulation → implementation → outcome, moderated by analytical, operational, and political capacities, and embedded in the good governance of evidence.

The framework is categorical rather than causal, and this study does not seek to test it statistically. Nevertheless, consistent with Yin's (2018) account of case study design, it is both possible and desirable to state in advance what the literature suggests one might expect, thereby enabling readers to assess whether the fieldwork confirmed, qualified, or disconfirmed those expectations. Accordingly, Table 1 presents four *orienting propositions* derived from the literature reviewed above, formulated before Stage 1 fieldwork commenced. These propositions serve as sensitizing devices that guide attention and specify what would

constitute disconfirming evidence rather than functioning as hypotheses to be tested against a sampling distribution. Each proposition is revisited in Section 5.

Table 1. Orienting propositions derived from the literature and formulated prior to fieldwork

Stage	Orienting proposition	Revisited in
Input	P1. Where metadata standards and walidata arrangements are not shared across agencies, the credibility of the evidence entering the cycle will be contested, and individual-level analytical capacity will yield limited returns (Parkhurst, 2017; Bernot et al., 2024).	Section 5.1
Formulation	P2. Where technocratic compilation and deliberative aggregation operate in parallel rather than in dialogue, evidence will be salient to neither planners nor citizens, and formulation will be dominated by political signals (Cairney, 2022; Anindito et al., 2022).	Sections 5.1, 5.3
Implementation	P3. Where no agency-wide analytical protocol exists, variations in evidence use will track individual training and leadership signalling rather than organizational design (Christensen, 2025; Howlett, 2015).	Section 5.2
Outcome	P4. Actors will associate EBP with improved performance in domains where input-stage data quality is strong; this association will be reported as a perception rather than demonstrated as an effect (Head, 2016; Aliyu et al., 2022).	Sections 5.2, 5.5

3. RESEARCH METHODS

3.1 Research design and epistemological stance

This study adopts a single-case, two-stage qualitative design (Yin, 2018). Its logic is abductive and interpretive: the four-stage framework provided an initial vocabulary; fieldwork revealed patterns that this vocabulary did not anticipate; and the framework was subsequently revised to accommodate them. The study neither tests hypotheses nor estimates effects, and it makes no claims of statistical generalizability. Its inferential warrant is analytic generalization —the extension of findings to theoretical propositions rather than to populations (Yin, 2018)—and it draws on Flyvbjerg's (2006) argument that a well-chosen single case can contribute to scientific development by generating context-dependent knowledge that large-N designs cannot access. We make this explicit because readers accustomed to the hypothetico-deductive format conventional in accounting and business research might otherwise interpret the absence of formal hypotheses as an omission rather than as a deliberate design choice.

3.2 Case selection

Kudus Regency, located in Central Java and home to approximately 870,000 residents, was selected as a *typical case* in Seawright and Gerring's (2008) sense: a case that is unexceptional on the dimensions defining the population of interest and therefore permits the investigation of the mechanism operating within that population. The population of interest is not "Indonesian local government" in general, but rather a narrower and more defensible set: mid-sized Javanese regencies that have formally adopted the post-2019 EBP architecture but have not yet institutionalized it in their routine practice. Kudus issued Regent Regulation No. 41 of 2021 on the Implementation of One Regional Data, established Bappeda and Diskominfo institutions, and conducted an active Musrenbang process. Formal adoption is therefore complete, while the gap between adoption and practice remains open, which is the phenomenon under study.

Formal adoption alone is a weak basis for a typicality claim, as it says nothing about where the case sits relative to comparable jurisdictions. To make the claim auditable rather than assertive, Table 2 positions Kudus against the population of Central Java regencies on five

publicly reported indicators. The criteria were specified before case selection and applied to all Central Java regencies that had issued a regional One Data regulation by the end of 2021.

Table 2. Case-selection criteria and the position of Kudus Regency relative to Central Java regencies

Criterion	Rationale for inclusion	Kudus position	Source
Population size	Mid-sized regencies face capacity constraints, unlike metropolitan cities or very small regencies.	below	BPS, 2022
Regional fiscal capacity index	Fiscal capacity affects the ability to fund analytical units and data infrastructure.	0.56 (middle quintile)	Ministry of Finance, 2021
Bureaucratic Reform Index (Indeks RB)	Proxies organizational maturity within which evidence use is embedded.	BB (Good)	KemenPAN-RB, 2021
E-Government Index (Indeks SPBE)	Proxies the technical substrate on which the Satu Data implementation depends.	3.10 (upper-middle category among Central Java regencies)	Kominfo, 2021
Formal adoption of regional One Data regulation	Defines the population of interest: adoption complete, institutionalization incomplete.	Regent Regulation No. 41 of 2021 — adopted	Kudus Regency Government (2021)

Kudus was positioned within the expected range of Central Java regencies: population of approximately 870,000 (BPS, 2022), moderate fiscal capacity (Ministry of Finance, 2021), Bureaucratic Reform Index in the "good" category (KemenPAN-RB, 2021), and SPBE implementation in the upper-middle category among Central Java regencies (Kominfo, 2021). Where regency-level comparative rankings were unavailable, the closest published provincial comparison was used.

According to these criteria, Kudus is unexceptional rather than exemplary, as required by typical-case logic. This does not imply that Kudus is representative of Indonesia as a whole; the boundary conditions for transfer are specified in Section 5.6.

3.3 Sampling strategy and informants

Sixteen in-depth, semi-structured interviews were conducted with informants purposively selected to represent the principal actors in the local EBP system in the study area. Table 3 presents the characteristics of the sample.

Table 3. Informant profile, Stage 1 (n = 16)

Code	Institution and role	n	Also participated in
BAP-01 to BAP-04	Bappeda: Head and senior planners (planning, statistics and evaluation functions)	4	FGD 1–3; workshops 1–2
OPD-01 to OPD-07	communication and informatics/data governance	7	FGD 1–3; workshops 1–2
DKI-01, DKI-02	Diskominfo walidata (data-custodian) team	2	FGD 1, 3; workshops 1–2
AKD-01, AKD-02	Academics, local universities	2	FGD 2; workshop 2
CSO-01	Civil-society representative	1	FGD 3; workshop 2

Seventh sector: communication and informatics/data governance. Academic informants participated in FGD 2, while the civil society informant participated in FGD 3 and workshop 2.

The interview guides were structured around the four stages of the analytical framework and explored the perceptions of evidence demand, supply, and use, as well as barriers. The informants were identified throughout the study using alphanumeric codes.

The seven OPDs were not selected opportunistically. The sample was determined by applying three criteria sequentially. First, the sector had to be responsible for a mandatory basic-service function (*urusan wajib pelayanan dasar*) under Law No. 23 of 2014 and subject to minimum service standard obligations, as routine outcome data are legally required for the health, education, public works, and social affairs sectors. Second, the sector had to have a designated data-governance role under Regent Regulation No. 41 of 2021, which established planning and statistical functions as producers and custodians of the regency's data holdings. Third, among the remaining candidates, priority was given to sectors with the largest shares of regency expenditure and the highest volumes of routine administrative data, on the grounds that the costs of inadequate evidence are greatest in these sectors.

Sectors such as agriculture, the environment, cooperatives, and small enterprises were therefore excluded despite their relevance to public services. This exclusion narrows the empirical base, which is a limitation that readers should consider. Sectors with extensive field-level data collection but limited statistical infrastructure, such as agriculture, and those that rely on externally sourced technical data, such as the environment, may exhibit input-stage dynamics that differ from those reported in Section 4.1. Accordingly, the findings are presented as applying primarily to mandatory basic service sectors, and the model in Section 5.4 should be applied to other sectors as a diagnostic tool rather than as an established description.

3.4 Data collection

Stage 1 (April–September 2022) used three methods. In addition to the 16 interviews, three focus group discussions (FGDs) were conducted, one for each of the input, formulation, and implementation/outcome stages. The participants included representatives from Bappeda, OPDs, Diskominfo, academia, and civil society. FGDs were selected to identify inter-organizational dynamics, including disagreements over data ownership and analytical standards, which individual interviews may tend to obscure. As noted in Section 5.5, the FGD findings therefore provide an important, though partial, means of assessing the limitations of self-reported data. Third, a documentary analysis examined regulatory instruments (Government Regulation No. 18 of 2016; Permendagri No. 86 of 2017; Presidential Regulation No. 39 of 2019; Regent Regulation No. 41 of 2021), the regency's RPJMD 2018–2023, OPD performance reports, and Musrenbang minutes from 2020 to 2022.

Stage 2 (October 2022–February 2023) tested the model developed in Stage 1 through two participatory validation workshops involving Stage 1 informants and invited external experts (n = 22). The participants reviewed the draft model, assessed each stage based on their institutional experiences, and identified potential revisions to the model. The workshop proceedings were transcribed and incorporated as a third tier of the data collection. *The workshops led to refinements of the model, including an enhanced role for coordination mechanisms – particularly the 3C principle – and clarification of the feedback relationship between outcome evaluation and future inputs. The four-stage structure (input, formulation, implementation, and outcome) was unchanged.*

3.5 Sampling adequacy and the saturation protocol

Sample size in qualitative research requires justification rather than assertion. Saturation was operationalized here in Saunders et al.'s (2018) *a priori thematic* sense: the point at which every cell of the four-stage × three-capacity coding matrix was populated by at least three independent informants, and at which two consecutive interviews generated no new first-order codes. A running code inventory was maintained after each interview, consistent with Guest, Bunce and Johnson's (2006) demonstration that the great majority of codes in a

relatively homogeneous purposive sample typically emerge within the first twelve interviews.

The last new first-order code emerged after the fourteenth interview. The final three interviews yielded two additional refinements, but no new first-order codes were identified. One matrix cell concerning political capacity at the outcome stage included fewer than three participants. This limitation is acknowledged and treated as partial coverage, rather than full saturation.

In this protocol, 16 interviews were considered adequate for addressing the a priori thematic categories derived from the framework. However, they are *not* claimed to have achieved theoretical saturation of the inductively derived 3C construct, which emerged late in the coding process and would require a sample designed specifically to examine it. This asymmetry is acknowledged as a limitation in Section 5.5 rather than subsumed under a general claim of sufficiency.

3.6 Data analysis and the operationalization of outcome constructs

All audio material was transcribed verbatim and imported into NVivo 12 for thematic analysis, following Braun and Clarke (2006), with the coding matrix and display conventions drawn from Miles, Huberman and Saldaña (2014). Coding proceeded in two cycles. The first applied a deductive coding structure derived from the analytical framework (input, formulation, implementation, outcome × analytical, operational, political). The second was inductive, identifying emergent themes—most notably, the *3C principle* (cooperation, communication, coordination), which informants raised repeatedly as a meso-level mechanism linking individual analytical work to inter-organizational outcomes. Two researchers coded the material independently and reconciled discrepancies through discussion, achieving an inter-coder agreement of $\kappa = 0.81$. A theme-by-informant matrix was produced from the coded corpus and is presented in Table 5.

The four outcome constructs warrant particular care, as they are the point at which perceptual data are most readily mistaken for performance measurement. Table 4 outlines how each construct was defined and examined in the interviews, the independent material against which it was triangulated, and its epistemic status.

Table 4. Operational definitions and epistemic status of the four outcome constructs

Construct	Working definition used in coding	Interview probe (abbreviated)	Triangulated against
Policy accuracy	Correspondence between the problem specified in a program's planning document and the problem evidenced in baseline data for the same population.	"How often do the targets in the RKPD match what your own data show about the problem?"	RPJMD 2018–2023; RKPD documents; OPD performance reports
Response speed	Elapsed time between the documented identification of a problem and formal launch of a programme addressing it, judged relative to the period before the One Data regulation.	"Compared with before 2021, how long does it take for a problem visible in data to become a programme?"	Musrenbang minutes 2020–2022; programme launch dates in OPD reports
Service affordability	Whether the programme design reduced the direct costs borne by intended beneficiaries.	"Did the way this programme was designed change what beneficiaries have to pay or travel?"	Programme design documents; budget allocations
Programme sustainability	Whether a program continued to be funded and delivered beyond a single budget cycle.	"Which of these programmes survived the next budget cycle, and why?"	Regency budget documents 2020–2022

Two properties of these constructs should be considered throughout Section 4. First, all four are judgements made by officials about domains for which they are themselves accountable; they measure *perceived* performance and are subject to the self-presentational pressures discussed in Section 5.5. Second, although each was triangulated against documentary material, that material corroborates the existence and timing of programs rather than the quality of the judgments about them. The analysis treats the four constructs accordingly, and the article does not convert them into performance measures.

3.7 Trustworthiness

Rigor was assessed using Lincoln and Guba's (1985) four criteria. *Credibility* was established through triangulation across interviews, focus group discussions (FGDs), and documents, as well as member checking with five Stage 1 informants who reviewed the draft analysis. *Dependability* was supported by a two-cycle coding protocol and inter-coder agreement (κ = 0.81). *Confirmability* was supported by an audit trail linking every claim in Section 4 to the coded extracts and, where applicable, the documentary sources. *Transferability* is addressed not by claiming generalizability, but by providing contextual details in Sections 3.2 and 5.6 that enable readers to assess which elements may be applicable elsewhere.

3.8 Ethics, reflexivity and limitations

This study was conducted in accordance with the ethical protocols of the host institution. Written informed consent was obtained, and the informants were assigned alphanumeric codes. Four limitations warrant explicit acknowledgment; each is carried forward for consideration rather than resolved herein.

First, the study employed a single-case design. Its findings were analytically rather than statistically generalizable. Section 5.6 specifies the contextual variables that determine the conditions under which the findings may be transferred.

Second, and most consequentially, the evidence from non-state actors is limited. Civil society representation in the interview sample consisted of a single informant (CSO-01), supplemented by civil society participation in the FGDs and Stage 2 workshops. This limitation directly affects the credibility of one specific claim: the finding reported in Section 4.2 that participation in Musrenbang is procedural rather than substantive. The claim is based on convergent but limited evidence—one civil society interview, civil society participation in *all three* FGDs, and documentary analysis of Musrenbang minutes from 2020 to 2022—and, critically, on accounts from the officials who administer the forum rather than from the citizens who attend it. Officials may have their own reasons for characterizing deliberation in particular ways, and the direction of any such bias is not clear a priori. We therefore present this finding as a proposition consistent with the broader literature (Anindito et al., 2022; Sutiyo & Maharjan, 2017) and with the partial evidence available to us, rather than as an established characterization of citizens' experience of Musrenbang, which our data cannot establish. A citizen survey or observational study of Musrenbang sessions would be necessary to support this stronger claim; we do not make it.

Third, the reliance on self-reports, as no coordination episodes were directly observed. The implications for the construct validity of the 3C principle are examined in Section 5.5.

Fourth, the sectoral scope is also important. As set out in Section 3.3, the OPD sample is limited to the mandatory basic service and data governance sectors.

4. FINDINGS

This section presents the findings across the framework's four stages. Two conventions apply throughout. First, descriptive claims are quantified wherever the data permit: statements such as "several OPDs" are replaced with counts out of the seven OPDs in the sample, enabling readers to independently assess the evidential basis for each claim. Second, illustrative

verbatim extracts, accompanied by informant codes, are provided for each stage, allowing readers to evaluate the interpretations presented here against the evidence on which they are based. Table 5 presents the distribution of the principal themes across informant groups.

Table 5. Distribution of principal themes across informant groups (number of informants coded at least once at each node; n = 16)

Theme	Bappeda (4)	OPD (7)	Diskominfo (2)	Academic (2)	CSO (1)	Total
Inconsistent metadata definitions across agencies	3	5	2	1	0	11
No dedicated analytical unit or post in the agency	1	5	1	0	0	7
Portal richer in administrative than in outcome data	2	6	2	1	0	11
Low external use of the open-data portal	1	4	2	2	1	10
3C invoked as the mechanism resolving evidence–politics conflict	3	4	1	2	1	11
Musrenbang as procedural rather than substantive deliberation	3	5	1	1	1	11
No agreed analytical protocol for programme data	2	6	1	0	0	9
Leadership signalling as a driver of analysis quality	3	4	1	0	0	8
Weak participation in implementation monitoring	1	5	1	1	1	9
EBP associated with better decisions where data quality is strong	3	6	2	2	1	14

4.1 Input stage: data infrastructure and analytical capacity

Informants consistently identified the input stage as the foundation of the EBP cycle and the area with the most significant capacity deficits. Three categories of data—baseline, basic, and technical—were recognized as essential, with Bappeda and Diskominfo holding formal authority and OPDs serving as data producers. The implementation of the Regency's One Regional Data regulation has established a formal *validata* function and an inter-OPD data forum that meets quarterly, as corroborated by the forum minutes reviewed during the documentary analysis. Nevertheless, four operational constraints recurred across interviews and focus group discussions (FGDs).

First, metadata standards are inconsistent across agencies. Informants from four of the seven OPDs, along with both Diskominfo informants, described instances in which the same indicator—most frequently, the poverty headcount—was measured using different definitions across agencies. This mirrors the cross-agency interoperability challenges identified nationally by Bernot et al. (2024) at the regency level.

> *"Ketika kami membandingkan data kemiskinan dari beberapa OPD, ternyata definisi dan periode pengambilan datanya tidak selalu sama. Setelah dibahas dalam forum data, kami menyepakati indikator yang digunakan bersama agar angka yang masuk dalam perencanaan tidak berbeda." (BAP-02)*

Second, analytical capacity was limited and unevenly distributed. Of the seven OPDs in the sample, four had no unit or designated position whose primary responsibility was data analysis. In the remaining cases, this function was assigned to a single planner alongside other

duties. In four cases, data collection responsibilities rested with junior planners who reported having received no formal training in statistical analyses.

> *"Di OPD kami belum ada analis khusus. Biasanya pengolahan data dilakukan oleh staf perencanaan yang juga mengerjakan tugas lain. Kemampuan analisis lebih banyak berdasarkan pengalaman dan pelatihan yang pernah diikuti masing-masing pegawai." (OPD-04)*

Third, although the Diskominfo team has established a regency data portal, its contents are unevenly distributed: administrative data (expenditure and beneficiary counts) are documented far more comprehensively than outcome data (service quality and program effectiveness). This asymmetry is significant for the framework, as the input stage provides information for monitoring compliance but not for evaluating outcomes. Consequently, it constrains the feedback loop from outcomes to inputs, as illustrated in Figure 1.
Fourth, civil society and academic informants reported that external actors do not use the open-data portal on a routine basis. According to Parkhurst's (2017) criteria, this reflects a legitimacy deficit rather than a technical one: credible and salient evidence for officials but remains invisible to those affected fails the representativeness test that distinguishes well-governed evidence from evidence that is merely available. Given the limited non-state sample (Section 3.8), this observation is presented as consistent with, rather than demonstrative of, low external use.

> *"Portal data daerah sudah menyediakan banyak informasi, tetapi pemanfaatannya oleh pihak luar masih terbatas. Akademisi biasanya tetap meminta data secara langsung karena belum semua data analisis atau data hasil tersedia dalam bentuk yang mudah digunakan." (AKD-01)*

4.2 Formulation stage: technocratic analysis and deliberative engagement

Policy formulation in Kudus is organized around the annual planning cycle. OPDs submit proposals based on sectoral data, which Bappeda integrates through technical compilation. The Musrenbang process then consolidates community priorities across subdistrict, regency, and provincial levels. Informants characterized the process as both technocratic and deliberative—a configuration consistent with the view that EBP cannot be separated from the political process (Davies et al., 2000; Cairney, 2022).
Two findings are particularly noteworthy in this regard. First, when conflicts arose between technical evidence and, for example, OPD analyses recommending budget reallocation—and political signals conveyed through Musrenbang or by elected officials- informants identified the *3C principle*—cooperation, communication, and coordination among the affected agencies—as a resolution mechanism. Four of the 16 informants used the term spontaneously. Rather than displacing evidence, the principle, as described, operationalizes Parkhurst's (2017) good governance of evidence by channeling technical findings through deliberative arrangements rather than circumventing them, as is often the case.

> *"Pernah ada usulan program yang berdasarkan data belum menjadi prioritas, tetapi muncul kuat dalam Musrenbang. Kami tidak langsung menolak, tetapi berkomunikasi dengan OPD terkait dan menyamakan persepsi. Setelah koordinasi, keputusan akhirnya tetap mempertimbangkan data sekaligus kebutuhan masyarakat." (BAP-01)*
> *"Koordinasi sudah dilakukan, tetapi dalam beberapa kondisi keputusan akhirnya tetap mengikuti arahan yang telah ditetapkan sebelumnya. Data menjadi bahan pertimbangan, namun tidak selalu menjadi faktor penentu karena ada pertimbangan lain di tingkat kebijakan." (OPD-02)*

Second, informants characterized Musrenbang as involving procedural rather than substantive deliberation. Citizen attendance was reportedly consistent, but influence over the final priority list was filtered through Bappeda's compilation process in ways that four informants described as opaque. This finding aligns with broader critiques of the forum

(Anindito et al., 2022). However, it should be interpreted with caution, as outlined in Section 3.8: it rests principally on the accounts of officials who administer the forum, supplemented by one civil society interview and documentary analysis of Musrenbang minutes, and does not reflect citizens' own experiences of participation.

> *"Masyarakat memang menyampaikan banyak usulan dalam Musrenbang, tetapi setelah masuk tahap penyusunan prioritas, tidak semua usulan terlihat kembali. Kami memahami ada keterbatasan anggaran, namun mekanisme penyaringan usulan belum selalu diketahui oleh masyarakat." (CSO-01)*

4.3 Implementation stage: bureaucratic capacity and leadership

The implementation phase revealed the most significant gap between formal procedures and operational realities. Formally, programs are translated into OPD work plans accompanied by performance indicators, in accordance with Permendagri No. 86 of 2017. In practice, the data highlighted three principal challenges.

First, no agreed-upon standard analytical method exists for processing program data. Among the seven OPDs, four reported relying solely on descriptive statistics, four reported using simple trend analyses, and four reported commissioning external academic consultations. The selection of methods was described as being determined by the individual analyst's training rather than an agency-wide protocol.

> *"Metode analisis biasanya kami sesuaikan dengan kebutuhan program dan kemampuan staf yang tersedia. Kalau datanya sederhana, kami menggunakan perbandingan tren, tetapi untuk analisis yang lebih mendalam, kadang kami membutuhkan bantuan pihak lain." (OPD-06)*

Second, leadership signaling is a significant factor. Informants in OPDs whose senior leaders had explicitly endorsed data-driven management reported substantially more advanced analytical practices than those in organizations without such endorsement. This pattern is consistent with Christensen's (2025) finding that organizational design and leadership independently influence evidence uptake, as well as with the demand-side argument presented by Howlett (2015).

> *"Ketika pimpinan mulai meminta setiap usulan disertai data pendukung, staf menjadi lebih terbiasa melihat data sebelum menyusun program. Sebelumnya, beberapa usulan lebih banyak berdasarkan pengalaman tahun sebelumnya." (OPD-03)*

Third, public participation in implementation monitoring was reported to be weaker than at the Musrenbang stage, despite being formally provided for in the regulations. Four informants raised this point.

4.4 Outcome stage: perceived indicators of local-government performance

The informants identified the four indicators defined in Table 4: policy accuracy, response speed, service affordability, and program sustainability. Based on these criteria, EBP was perceived to have improved decision quality in domains with strong data quality—particularly health and education—and to have made marginal contributions in domains where data quality remained weak, particularly in labor and the informal economy. This pattern is consistent with the expectations of Proposition P4.

> *"Di bidang kesehatan, data cakupan layanan membantu kami menentukan kelompok sasaran dengan lebih tepat. Namun, untuk sektor yang datanya belum lengkap, seperti pekerja informal, kami masih kesulitan memastikan apakah program tersebut sudah menjawab kebutuhan yang sebenarnya." (OPD-01)*

The epistemic status of this pattern should be explicitly stated rather than assumed. These findings reflect governance perceptions rather than causal estimates. The research design does not establish that EBP produced performance gains; rather, it shows that actors closely involved in the policy process associated EBP investments with higher-quality decisions, particularly when they considered the 3C principle operational. This association may reflect genuine improvements; alternatively, it may reflect the greater visibility of improvements in domains with better measurements or officials' tendency to attribute favorable outcomes to reforms they have implemented. The present data cannot distinguish between these two interpretations. The data support the more limited claim that the perceived returns to EBP vary across sectors in a manner that corresponds to input-stage data quality. This finding is consistent with broader cross-national evidence indicating that local government effectiveness in decentralized systems depends on the combined operation of analytical, operational, and political capacities (Aliyu et al., 2022; Hummel & Kusumasari, 2025).

5. DISCUSSION

5.1 EBP as a politically governed cycle, not a linear pipeline

The Kudus findings provide empirical support for the theoretical shift from EBP to the good governance of evidence (Parkhurst, 2017) and directly inform propositions P1 and P2. At each of the four stages, technical evidence work was embedded in inter-organizational relationships, leadership signals, and deliberative arenas. The findings support P1: contested metadata definitions were reported to undermine the credibility of evidence at the point of entry, while informants in agencies with trained analysts described the training as offering limited value when the underlying data remained disputed. P2 receives more qualified support. As anticipated, technocratic compilation and deliberative aggregation operated substantially in parallel. However, the 3C principle occupied precisely the interface where the proposition predicted a void. This finding was not anticipated by the literature and required a revision of the framework rather than its mere confirmation.

The 3C principle, as described, is not an alternative to EBP but rather a vernacular articulation of what Cairney (2022) terms evidence use under decentered governance: evidence shapes decisions when it is negotiated with political actors rather than when it bypasses them. Viewed in light of the collaborative-governance literature, the mechanism is recognizable: cooperation, communication, and coordination closely correspond to the face-to-face dialogue, trust-building, and shared understanding that Ansell and Gash (2008) identify as the internal dynamics of collaboration, as well as the principled engagement and shared motivation components of Emerson et al.'s (2012) framework. What is distinctive in the Kudus case is the locus: these dynamics operate *among agencies within a single local government* rather than between state and non-state actors, and they are invoked as a bureaucratic norm rather than established through a designed forum. This relocation constitutes the conceptual contribution and also gives rise to the validity problem examined in Section 5.5. The feedback dynamics implied by this account connect to the broader literature on policy learning (Dunlop & Radaelli, 2018), which provides a vocabulary for the evaluation loop posited in Figure 1 but shown by the Kudus data to be weakly instantiated.

The practical implication is clear: training that equips planners with analytical skills alone, without a corresponding investment in inter-organizational relationships through which analysis is communicated, is unlikely to result in changed decisions.

5.2 Capacity complementarities and the locus of intervention

Consistent with the policy capacity literature (Wu et al., 2015; Mukherjee et al., 2021; Aliyu et al., 2022), Kudus demonstrates more pronounced deficits in certain dimensions of the capacity matrix than others. The operational capacity at the systemic level—particularly data interoperability, agreed metadata, and walidata staffing—emerges as the most significant

constraint. This is followed by analytical capacity at the individual level (training for OPD planners) and political capacity at the organizational level (leadership signalling by regents and OPD heads). Proposition P3 is supported: in the absence of agency-wide analytical protocols, variation in evidence use was associated with individual training and leadership signals rather than organizational design.

This implies that capacity-building investments should be sequenced rather than dispersed: addressing data interoperability first establishes the foundation on which subsequent investments in analytical training can generate returns. This reflects the local government analog of Howlett and Ramesh's (2016) "Achilles' heels" argument, according to which marginal investment in the weakest capacity yields the greatest gains. Two caveats apply. First, the identification of operational capacity as the binding constraint is specific to Kudus; in a regency with weaker human capital or more limited connectivity, the ordering could differ, as discussed in Section 5.6. Second, the sequencing claim is a design inference drawn from the observed pattern of deficits rather than a demonstrated result; it has not been tested against a case in which the sequence was reversed.

5.3 Articulating Satu Data Indonesia with deliberative arenas

The *Satu Data Indonesia* architecture (Bernot et al., 2024) and Musrenbang tradition (Anindito et al., 2022) are typically examined in separate bodies of literature; the former is technocratic, while the latter is deliberative. The Kudus case suggests that effectively integrating these approaches is the central operational challenge for evidence-based policymaking (EBP) in Indonesian local government. When data were presented in an accessible format during Musrenbang—as in the example provided by informants of sub-district-level service-coverage maps—deliberations were described as more substantive and policy choices as more accurate. When data and deliberation operated in parallel rather than in dialogue, informants described both processes as losing effectiveness: data became administratively disconnected, while deliberation became procedurally hollow, losing its purpose.

Designing this articulation is, in effect, a translation problem: converting walidata outputs into formats that citizens and elected officials can use in real time. The comparative literature on digitalized local administration supports treating this as an organizational rather than a technological task: Hirsch et al. (2022) find that the effectiveness of digital communication instruments in local authorities depends on pre-existing routines rather than on the instruments themselves, while Höchtl et al. (2016) make a parallel argument regarding the integration of large-scale data into the policy cycle.

5.4 The proposed evidence-based policy formulation model

Based on these findings, this article proposes a four-stage EBP formulation model adapted to the Indonesian local government (Figure 2). The model conceptualizes EBP as a cycle rather than a pipeline, specifies the capacity requirements at each stage, and identifies the deliberative-translational interfaces at which technical evidence and political processes must converge. The four stages are as follows: (i) *Input* — institutionalized walidata, agreed metadata, and an accessible open-data portal covering both outcome and administrative data; (ii) *Formulation* — technical compilation by Bappeda, linked to a substantive Musrenbang process through citizen-legible visual data products; (iii) *Implementation* — agreed analytical protocols, leadership signalling, and embedded researcher–practitioner relationships; and (iv) *Outcome* — the four indicators listed in Table 4, with structured feedback into the input stage.

Figure 2. Proposed evidence-based policy formulation model for Indonesian local government

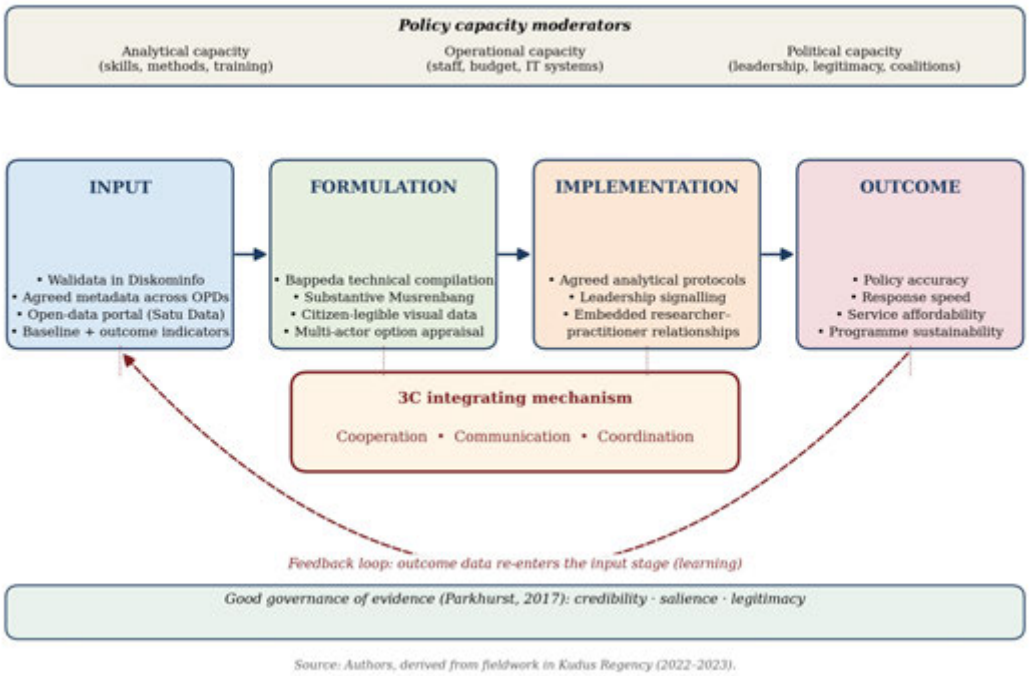


Figure 2. The proposed evidence-based policy formulation model for the Indonesian local government integrates the Satu Data Indonesia data architecture (input) with deliberative arenas (Musrenbang, public consultation) and the 3C principle (cooperation, communication, coordination) as the meso-level coordinating mechanism.

Three features distinguish this model from generic evidence-based policy (EBP) frameworks, such as the linear knowledge-to-policy pipeline and the OECD (2020) capacity checklist. These features merit explicit attention as they underpin the value of the contribution.

First, the model is architecturally specific. Its input stage is defined not by "good data" in the abstract but by the concrete institutional components of *Satu Data Indonesia*: the walidata designation, a shared metadata dictionary, a regency data portal, and the *forum data* mechanism established under Regent Regulation No. 41 of 2021. Whereas a generic framework instructs a local government to strengthen its evidence base, this model specifies which regulatory instrument, office, and artifact must be changed.

Second, the model is deliberately anchored in deliberation. The formulation stage does not treat public consultation as an optional legitimizing addition; rather, it identifies Musrenbang as the forum in which evidence must be made accessible to citizens and defines the translation of walidata outputs into visual, sub-district-level data products as the design task at this interface. Generic EBP and deliberative-planning frameworks tend to treat one another as contextual considerations; here, they constitute two stages of a single cycle.

Third, the model identifies a meso-level integrating mechanism. The 3C principle fills the gap left by generic frameworks—the space between an individual analyst's output and an inter-organizational decision. Frameworks in the policy-capacity tradition specify the competencies an organization requires, whereas frameworks in the EBP tradition specify the qualities evidence must possess. However, neither specifies who must communicate with whom, about which artifact, and at what stage. This model does so. This specificity makes it useful as a diagnostic instrument and limits its portability—a limitation discussed in Section 5.6.

5.5 Alternative interpretations and threats to the construct validity of the 3C principle

The 3C principle bears substantial theoretical weight in this study, yet it is also the construct for which the evidence is least robust. This section presents arguments for and against it with equal care.

The central threat is the management of impressions. "Kerja sama, komunikasi, koordinasi" is a phrase with strong normative resonance in Indonesian bureaucratic discourse; it is the vocabulary in which officials are trained to describe good administrative practice. Informants may therefore have invoked it as an account of how coordination *ought* to function rather than as a description of how a particular decision was actually reached. Because the construct emerged inductively from interviews and is evidenced by those same interviews, convergence across informants—ordinarily a mark of robustness—is partly uninformative in this instance: a shared professional vocabulary can produce convergent accounts regardless of whether it produces convergent behavior. This is the reflexive problem Alvesson (2003) identifies when treating interview talk as a transparent window onto organizational practice, compounded by the fact that no coordination episode was directly observed.

A second interpretation is that 3C is real but epiphenomenal: agencies coordinate because a shared budget ceiling and a common political principal compel them to do so, and "3C" is the label applied retrospectively to a compliance behavior that would occur regardless of the label. In this interpretation, the mechanism has descriptive validity but no explanatory value, and interventions designed to strengthen 3C would change the terminology without changing the coordination.

Three pieces of partial counter-evidence weigh against the strongest versions of these alternatives, although none is decisive. Disagreements over data ownership emerged openly in the FGD setting, suggesting that informants did not uniformly present a harmonious front. Documentary analysis confirms that the quarterly *forum data* meetings took place, establishing at least the infrastructure for coordination. Informants also linked 3C to specific, dateable episodes rather than describing it only in general terms. *One disconfirming instance was identified in which coordination occurred but did not affect the final decision, as broader political considerations remained decisive. This case is retained to demonstrate that 3C functions as a facilitating mechanism, rather than a guaranteed substitute for political judgment.*

Therefore, we present 3C as a candidate meso-level mechanism that warrants observational and process-tracing research, rather than as a validated construct. Readers should note an asymmetry in the evidence base of the article: the capacity deficits reported in Sections 4.1 and 4.3 are corroborated by documentary evidence and disagreement among informant groups, whereas the 3C mechanism relies more heavily on convergent self-report—a configuration in which impression management is particularly difficult to detect. Resolving this question would require direct observation of forum data and Musrenbang sessions, process-tracing of a small number of decisions from analysis through allocation, and examination of whether documented cross-agency data-sharing agreements were implemented in practice.

5.6 Boundary conditions and the limits of transfer

The model was developed in a single, mid-sized Javanese regency. While acknowledging that contexts in eastern Indonesia differ is necessary, it is not sufficient; practitioners and researchers need to know *which* contextual variables condition the model and *how* those variables would affect its application. Table 6 provides this specification in the form of adaptation questions rather than a validated moderation model table. It follows that the model should be characterized as a *transferable diagnostic template* rather than a generalizable model of local EBP. What is transferable is the sequence of diagnostic questions— which stage constitutes the binding constraint in a given context, at which interface evidence and deliberation fail to converge, and which capacity level is deficient—along with the identification of these interfaces as the areas where design efforts should be concentrated. What is not transferable is the substantive finding that systemic operational capacity is the

binding constraint; this claim pertains to Kudus and, at most, to jurisdictions resembling it along the dimensions presented in Table 6.

Table 6. Contextual moderators conditioning the application of the proposed model

Contextual variable	How it varies across Indonesian jurisdictions	Expected consequence for the model
Fiscal capacity	The own-source revenue and fiscal capacity index varies by an order of magnitude between resource-rich, Javanese, and remote regencies.	When fiscal capacity is low, the input stage cannot be addressed by staffing and systems investment; the binding constraint shifts to intergovernmental data provision, and the sequencing advice of the model reverses.
Connectivity to national data infrastructure	Bandwidth, electricity reliability, and physical distance from Bappenas and BPS structures differ sharply between Java and eastern Indonesia.	Portal-centered input designs presuppose reliable connectivity. Where it is absent, the walidata function must be re-specified around periodic offline reconciliation rather than through continuous interoperability.
Administrative human-capital depth	The availability of staff with statistical training and the presence of universities able to supply analytical partnerships are concentrated in Java.	Embedded researcher–practitioner relationships in the implementation stage are not universally available and may require substitution by provincial or regional analytical services.
Social heterogeneity and the character of deliberation	Ethnic, religious, and customary (adat) heterogeneity alters who speaks in deliberative arenas and how authority is recognized.	Musrenbang cannot be assumed to be the only or the most legitimate deliberative arena; the formulation stage must be re-specified around whichever arenas actually carry legitimacy locally.
Political competitiveness	Electoral competitiveness varies and conditions responsiveness independently of the planning procedure (Fossati, 2016).	Leadership signalling, on which the implementation stage depends, is itself endogenous to political incentives; where competition is low, the political capacity component cannot be treated as a design variable.
Sectoral data maturity	Basic-service sectors have mandated indicators; agriculture, environment, and informal economy sectors do not.	The model applies most directly to mandatory basic-service sectors (see Section 3.3); elsewhere, the input stage requires prior construction of indicators rather than standardization of existing ones.

5.7 Implications for theory

This article contributes to four bodies of literature, in each case in a bounded manner. For the evidence-based policymaking (EBP) literature, it provides Indonesian empirical grounding for the post-Parkhurst shift toward the good governance of evidence and identifies a meso-level mechanism at the district scale that has not been examined in the predominantly national and international literature. For the policy-capacity literature, it demonstrates the empirical applicability of the Wu–Ramesh–Howlett framework at the regency level and identifies systemic operational capacity as the binding constraint in one mid-sized Indonesian local government. This finding complements, and is not intended to supersede, Newman et al.'s (2022) provincial-level analysis of West Java. For the collaborative-governance literature, it presents a case in which the dynamics theorized by Ansell and Gash (2008) and Emerson et al. (2012) for cross-sector partnerships appear instead as an intra-governmental bureaucratic norm. This raises the question of whether collaborative dynamics require deliberately designed forums or can be sustained through professional vocabulary—a question that, as Section 5.5 shows, this study cannot resolve. For the decentralization literature, it adds district-level detail to the broader argument (Shoesmith et al., 2020; Hummel & Kusumasari,

2025) that the developmental benefits of decentralization depend on local governance configurations rather than arising automatically.

6. CONCLUSION

Two decades after the launch of Indonesian decentralization, evidence-based policymaking at the local level remains a work-in-progress. Drawing on a two-stage qualitative case study of the Kudus Regency, this article describes evidence-based policymaking as a four-stage cycle—input, formulation, implementation, and outcome—moderated by analytical, operational, and political capacities and operating through both technocratic and deliberative arrangements.

The central observation of the article is not that EBP improves local-government performance; the research design cannot support such a claim, and no such claim is made. Rather, the article observes that the actors operating the system explained the gap between the formal adoption of EBP regulations and their practical effects in terms of the capacity to connect the *Satu Data Indonesia* data architecture with deliberative forums such as Musrenbang. Where informants described this connection as having been achieved, they attributed it to cooperation, communication, and coordination among the agencies concerned—the 3C principle. Whether this connection leads to the performance improvements that informants associate with it is a question that this research design cannot resolve and is the first item on the research agenda set out below.

Three implications for practice follow. These are presented as design propositions consistent with the evidence, rather than as demonstrated causal recommendations. First, governments seeking to strengthen EBP should consider sequencing capacity investments by prioritizing interoperable data infrastructure before analytical training, as the latter was reported to yield limited benefits where the former was contested. Second, deliberative reform of Musrenbang—particularly the development of citizen-accessible data products—addresses the interface at which the present study found evidence and deliberation most frequently failed to converge. Third, leadership signalling appears consequential: regents and agency heads who require data-supported justifications establish the demand-side conditions that enable analytical investment to be put to use (Howlett, 2015).

Four avenues for further research follow directly from the limitations acknowledged above. First, the 3C principle requires observational and process-tracing investigations—direct observation of forum data and Musrenbang sessions, as well as tracing specific decisions from analysis through to allocation—to distinguish a substantive coordination mechanism from a shared professional vocabulary (Section 5.5). Second, the model should be tested comparatively, particularly in eastern Indonesian regencies whose positions on the moderators in Table 6 differ markedly from that of Kudus. Third, mixed-methods designs that pair the qualitative depth of the present study with measured indicators of policy accuracy and service delivery would allow the perceptual outcome constructs in Table 4 to be assessed against performance data, transforming the model from a diagnostic template into a testable causal framework. Fourth, the diffusion of artificial-intelligence-supported analytics into Indonesian government data systems (Giest et al., 2025) raises new questions about how machine-generated evidence will reshape the input and formulation stages—questions that are better addressed before these technologies become entrenched than afterward.

ACKNOWLEDGMENTS

The authors thank Bappeda, the OPDs, and Diskominfo of Kudus Regency for their cooperation, as well as the academic and civil society informants who participated in interviews, focus group discussions (FGDs), and validation workshops. The authors also thank the anonymous reviewer, whose comments substantially enhanced the article's epistemic precision, methodological transparency, and reporting of evidence. All remaining errors and interpretations are the authors' own.

FUNDING AND CONFLICTS OF INTEREST

This study received n*o external funding*. The authors declare no conflict of interest.

REFERENCES

Aliyu, A., Soeparman, S., & Hartono, B. (2022). The impact of local government capacity on public service delivery: Lessons learned from decentralized Indonesia. *Economies*, 10(12), 323. https://doi.org/10.3390/economies10120323

Alvesson, M. (2003). Beyond neopositivists, romantics, and localists: A reflexive approach to interviews in organizational research. *Academy of Management Review*, 28(1), 13–33. https://doi.org/10.2307/30040687

Anindito, D. B., Sagala, S. A. H., & Tarigan, A. K. M. (2022). e-Musrenbang: A digital framework for local participatory planning at the community level. *Town Planning Review*, 93(2), 137–160. https://doi.org/10.3828/tpr.2021.21

Ansell, C., & Gash, A. (2008). Collaborative governance in theory and practice. *Journal of Public Administration Research and Theory*, 18(4), 543–571. https://doi.org/10.1093/jopart/mum032

Bernot, A., Tjondronegoro, D., Rifai, B., Hasan, R., Liew, A. W. C., Verhelst, T., & Tiwari, M. (2024). Institutional dimensions in open government data: A deep dive into Indonesia's Satu Data initiative and its implications for developing countries. *Public Performance & Management Review*, 47(5), 1156–1184. https://doi.org/10.1080/15309576.2024.2377609

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cairney, P. (2016). *The politics of evidence-based policy making*. London: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-51781-4

Cairney, P. (2022). The myth of 'evidence-based policymaking' in a decentred state. *Public Policy and Administration*, 37(1), 46–66. https://doi.org/10.1177/0952076720905016

Christensen, J. (2025). Evidence and policy-making: An organizational approach. *Public Administration Review*. Advance online publication. https://doi.org/10.1111/puar.70051

Davies, H. T. O., Nutley, S. M., & Smith, P. C. (Eds.). (2000). *What works? Evidence-based policy and practice in public service*. Bristol: Policy Press. https://doi.org/10.51952/9781847425119

Dunlop, C. A., & Radaelli, C. M. (2018). Does policy learning meet the standards of an analytical framework of the policy process? *Policy Studies Journal*, 46(S1), S48–S68. https://doi.org/10.1111/psj.12250

Emerson, K., Nabatchi, T., & Balogh, S. (2012). An integrative framework for collaborative governance. *Journal of Public Administration Research and Theory*, 22(1), 1–29. https://doi.org/10.1093/jopart/mur011

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(2), 219–245. https://doi.org/10.1177/1077800405284363

Fossati, D. (2016). Beyond "good governance": The multilevel politics of health insurance for the poor in Indonesia. *World Development*, 87, 291–306. https://doi.org/10.1016/j.worlddev.2016.06.020

Giest, S., McBride, K., Nikiforova, A., & Sikder, S. K. (2025). Digital and data-driven transformations in governance: A landscape review. *Data & Policy*, 7, e21. https://doi.org/10.1017/dap.2024.47

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59–82. https://doi.org/10.1177/1525822X05279903

Head, B. W. (2016). Toward more "evidence-informed" policy making? *Public Administration Review*, 76(3), 472–484. https://doi.org/10.1111/puar.12475

Hirsch, B., Schäfer, F.-S., Aristovnik, A., Kovač, P., & Ravšelj, D. (2022). The impact of digitalised communication on the effectiveness of local administrative authorities — findings from Central European countries in the COVID-19 crisis. *Journal of Business Economics*, 92(7), 1009–1042. https://doi.org/10.1007/s11573-022-01106-8

Höchtl, J., Parycek, P., & Schöllhammer, R. (2016). Big data in the policy cycle: Policy decision making in the digital era. *Journal of Organizational Computing and Electronic Commerce*, 26(1–2), 147–169. https://doi.org/10.1080/10919392.2015.1125187

Howlett, M. (2015). Policy analytical capacity: The supply and demand for policy analysis in government. *Policy and Society*, 34(3–4), 173–182. https://doi.org/10.1016/j.polsoc.2015.09.002

Howlett, M., & Ramesh, M. (2016). Achilles' heels of governance: Critical capacity deficits and their role in governance failures. *Regulation & Governance*, 10(4), 301–313. https://doi.org/10.1111/rego.12091

Hummel, D., & Kusumasari, B. (2025). Local government financial performance and decentralization in Indonesia. *State and Local Government Review*. Advance online publication. https://doi.org/10.1177/0160323X251364379

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: SAGE.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Thousand Oaks, CA: SAGE.

Mukherjee, I., Coban, M. K., & Bali, A. S. (2021). Policy capacities and effective policy design: A review. *Policy Sciences*, 54(2), 243–268. https://doi.org/10.1007/s11077-021-09420-8

Newman, J., Cherney, A., & Head, B. W. (2017). Policy capacity and evidence-based policy in the public service. *Public Management Review*, 19(2), 157–174. https://doi.org/10.1080/14719037.2016.1148191

Newman, J., Patmisari, E., & Widianingsih, I. (2022). Policy analytical capacity and "Eastern" styles of policy analysis: Evidence from West Java Province, Indonesia. *Policy Sciences*, 55(3), 469–485. https://doi.org/10.1007/s11077-022-09470-6

Nutley, S. M., Walter, I., & Davies, H. T. O. (2007). *Using evidence: How research can inform public services*. Bristol: Policy Press. https://doi.org/10.46692/9781847422323

Oliver, K., Innvar, S., Lorenc, T., Woodman, J., & Thomas, J. (2014). A systematic review of barriers to and facilitators of the use of evidence by policymakers. *BMC Health Services Research*, 14(1), 2. https://doi.org/10.1186/1472-6963-14-2

Parkhurst, J. (2017). *The politics of evidence: From evidence-based policy to the good governance of evidence*. London: Routledge. https://doi.org/10.4324/9781315675008

Saguin, K., Ramesh, M., & Howlett, M. (2018). Policy work and capacities in a developing country: Evidence from the Philippines. *Asia Pacific Journal of Public Administration*, 40(1), 1–22. https://doi.org/10.1080/23276665.2018.1436427

Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. *Public Administration*, 80(1), 1–22. https://doi.org/10.1111/1467-9299.00292

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, 52(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Seawright, J., & Gerring, J. (2008). Case selection techniques in case study research: A menu of qualitative and quantitative options. *Political Research Quarterly*, 61(2), 294–308. https://doi.org/10.1177/1065912907313077

Shoesmith, D., Franklin, N., & Hidayat, R. (2020). Decentralised governance in Indonesia's disadvantaged regions: A critique of the underperforming model of local governance in eastern Indonesia. *Journal of Current Southeast Asian Affairs*, 39(3), 359–380. https://doi.org/10.1177/1868103420905479

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: SAGE.

Sutiyo, & Maharjan, K. L. (2017). *Decentralization and rural development in Indonesia*. Singapore: Springer. https://doi.org/10.1007/978-981-10-3208-0

Sutmuller, P. M., & Setiono, I. (2011). *Diagnostic on evidence-based public policy formation under decentralisation*. Jakarta: AusAID Australia–Indonesia Partnership for Decentralisation.

Weible, C. M., & Cairney, P. (Eds.). (2021). *Practical lessons from policy theories*. Bristol: Bristol University Press. https://doi.org/10.51952/9781529215687

Wu, X., Ramesh, M., & Howlett, M. (2015). Policy capacity: A conceptual framework for understanding policy competences and capabilities. *Policy and Society*, 34(3–4), 165–171. https://doi.org/10.1016/j.polsoc.2015.09.001

Wu, X., Howlett, M., & Ramesh, M. (Eds.). (2018). *Policy capacity and governance: Assessing governmental competencies and capabilities in theory and practice*. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-54675-9

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Thousand Oaks, CA: SAGE.

Legal and regulatory sources

Government of the Republic of Indonesia. (1999). *Law No. 22 of 1999 on Regional Government*. Jakarta: State Secretariat.

Government of the Republic of Indonesia. (2014). *Law No. 23 of 2014 on Regional Government*. Jakarta: State Secretariat.

Government of the Republic of Indonesia. (2016). *Government Regulation No. 18 of 2016, concerning Regional Apparatus*. Jakarta: State Secretariat.

Government of the Republic of Indonesia. (2017). *Regulation of the Minister of Home Affairs No. 86 of 2017 concerning Procedures for Planning, Controlling, and Evaluating Regional Development*. Jakarta: Ministry of Home Affairs.

Government of the Republic of Indonesia. (2018). *Presidential Regulation No. 95 of 2018 concerning Electronic-Based Government Systems*. Jakarta: State Secretariat.

Government of the Republic of Indonesia. (2019). *Presidential Regulation No. 39 of 2019 concerning One Data Indonesia (Satu Data Indonesia)*. Jakarta: State Secretariat.

Kudus Regency Government, Indonesia. (2021). *Regent Regulation No. 41 of 2021 concerning the Implementation of Regional Data*. Kudus: Bappeda Kabupaten Kudus.

Organisation for Economic Co-operation and Development [OECD]. (2020). *Building capacity for evidence-informed policy-making: Lessons from country experiences*. OECD Public Governance Review. Paris: OECD Publishing. https://doi.org/10.1787/86331250-en